



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : <b>H04N 7/15</b></p>	<p><b>A1</b></p>	<p>(11) International Publication Number: <b>WO 99/63756</b> (43) International Publication Date: 9 December 1999 (09.12.99)</p>
<p>(21) International Application Number: PCT/IT98/00149 (22) International Filing Date: 4 June 1998 (04.06.98) (71)(72) Applicant and Inventor: TRINCA, Roberto [IT/IT]; Via Casetta Mattei, 151, I-00148 Roma (IT). (74) Agent: SARPI, Maurizio; Studio Ferrario, Via Collina, 36, I-00187 Roma (IT).</p>		<p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p>Published With international search report.</p>
<p>(54) Title: PROCESS FOR CARRYING OUT VIDEOCONFERENCES WITH THE SIMULTANEOUS INSERTION OF AUXILIARY INFORMATION AND FILMS WITH TELEVISION MODALITIES</p> <div data-bbox="467 1163 1078 1587"> </div> <p>(57) Abstract</p> <p>Process for carrying out videoconferences among remote and/or local users, comprising the following steps: link-up to a direction room (1) with a plurality of both remote and neighbour locations (2), which an audiovisual signal (AV) originates at; if necessary, conversion of the signal (AV) from each location (2), before its transfer to the direction room (1), so as to make it suitable to the type of transmission being utilised; reversion of the received signal (AV), if necessary, before its arrival at the direction room (1); selection of the signal or signals to use and send away to the attendants and the speaker respectively, by an input audio video matrix (MV1); addition of the contributions and the necessary audio and/or video effects, titles, soundtracks, graphs and so on, by a video mixer; selection of the processed audiovisual signals (AV) and sending thereof to the several remote locations (2), according to the role that the users play at that moment (attendants or speakers).</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

PROCESS FOR CARRYING OUT VIDEOCONFERENCES WITH THE SIMULTANEOUS INSERTION OF AUXILIARY  
INFORMATION AND FILMS WITH TELEVISION MODALITIES

DESCRIPTION

The present invention relates to the field of multimedia communications, and more particularly a process and apparatus therefor for videoconferences that provides link-ups among several attendants and  
5 with extremely variable characteristics and modalities, adaptable to any specific need of the user.

Currently, multiple user videoconference apparatus and techniques are known, and despite being based on different execution parameters, they make the choice of  
10 the image to be shown to the attendants on the grounds of the audio signal coming from the attendants themselves, which is technically called "audio presence".

In other words, the sound received by the  
15 microphone located at every equipped location gets to the centralised videoconference management device. This device shows all the attendants the image of the user that has generated the sound impulse. In such a way, all the attendants receive the image of the person  
20 that is speaking at that precise moment of time on their screens. It is therefore clear that if two or more users speak at the same time, the conference management device carries out image commutations on a continuous basis, causing considerable disruptions and  
25 chaos all along the course of the videoconference

itself.

Attention is also drawn to the fact that a user is allowed into a dialogue which has already started, even because of a background noise from his own environment, which could be completely independent from his will but is detected by the microphone located at his place.

Currently, in order to resolve such problems, it is necessary to turn off one's own microphone (but this risks turning an interesting debate into an endless monologue).

This type of automatic commutation caused by the audio presence, necessarily requires the presence of an interpreter next to each single attendant, in case of videoconferences that involve people speaking different languages.

Besides this, current technology does not always provide carrying out a link-up between different videocommunication systems. The apparatus which are currently being used in fact only allow file transmission and/or sharing just in case the link-up devices of the several attendants are made by the same manufacturer, in so doing drastically limiting the possibilities of employment of the system itself (file sharing, transmission and transfer, etc.).

A further problem of the prior art is given by the fact that the possibility of executing fadings among the images of the speakers that make their contributions along the way, and possible audio-video contributions, whether they be films, photographs, static images, graphs and so on, is ruled out.

A further disadvantage of the currently viable videoconferences, is given by the fact that it is not possible to superimpose titles, subtitles, abbreviations, speakers' names, musical themes and soundtracks, and all audio and video effects that can make of a "flat" and static videoconference a real television programme.

In this respect, it is useful to observe that said problems and drawbacks have not only got purely aesthetic consequences, but they also cause a rapid decrease in the level of attention of the attendants, which is an extremely important factor for the success of a conference of whatever type.

A first aim of the present invention is that of allowing the course of videoconferences (congresses, debates, presentations, lectures, etc.) with the utilization of audiovisual contributions such as films, slides, photographs, animated computer aided design, graphs, music and/or soundtracks etc.

A second aim of the present invention is that of guaranteeing an orderly and fluent course of a videoconference, thanks to the audio-visual commutations carried out by the operators of the direction room, and by the possible presence of a chairperson, who is meant to allow the user to personally take part in the debate, only at the most suitable moment.

A third aim of the present invention is that of giving the possibility to attend a conference even to Internet users. Furthermore, through a series of

procedures and suitable links, which will be analysed in detail in the foregoing, giving the possibility to any single spectator who is suitably equipped to directly enter and take part in the conference, contributing to it with his own image and his own audio (even if not originally scheduled).

A fourth aim of the present invention is that of guaranteeing compatibility between different videocommunication systems, utilising the most suitable interfaces and transforming the ensemble of the videoconference into many point-point links (user-direction) with personalised characteristics and communication protocols.

To this purpose attention is drawn to the fact that attendants, whether they be interactive or not, can be both remote and local and numberwise limitless.

These and other aims have been accomplished according to the invention, by proposing a process and an apparatus for the production and management of videoconferences, wherein audiovisual signals coming from a plurality of remote and/or neighbour locations, are acquired and elaborated by a direction room capable of dealing with and selecting both the audio and the video signal, adding audiovisual contributions like television effects, partial or total image superimposition, insertion of graphs, tables, films or soundtracks, audio commentaries, and so on.

According to the process and the apparatus which are herein described, it is also possible to provide a centralised interpretation service, discriminating on

the audio supplied by the users as a function of their language.

A better understanding of the present invention will be gained thanks to the following detailed description with reference to the appended drawings, which schematically illustrate a preferred embodiment of the invention.

In the drawings:

Fig. 1 schematically illustrates the parts making up the direction room according to the present invention;

Fig. 2 is a scheme illustrating the modalities and possibilities of link-up between the direction room and remote and/or neighbour users, by use of telephone lines, via satellite, via Internet, and so on.

With reference to the abovementioned figures, the process object of the present invention comprises the following stages:

-link-up in a direction room 1 with a plurality of remote and/or neighbour locations 2, which generate an audio video signal AV;

-conversion, if necessary, of the audiovisual signal AV from every location, before its transfer from the place where it is generated to that where the direction room 1 is located, to adapt it to the type of connection and transmission which is employed;

-reconversion of the received signal, if necessary, to an audio-video format, before its entrance to direction room 1;

-selection of the signal/s to be used and sent,

respectively, to the attendants and the speakers by an entrance audio-video matrix MV1;

-addition of the necessary audiovideo contributions and effects, as well as of titles, soundtracks, commentaries, graphs, and so on, by mixer  
5 video MIX or computer with analogous functions:

-selections of the processed audiovideo signals and their forwarding to the several remote locations 2, as a function of the fact that at that moment they are  
10 attendants or speakers.

According to a particular aspect of the process described above, while the attendants receive the audio video signal from the speaker, the latter will be capable of receiving a different audio signal which has  
15 been selected by the direction room.

For example, the speaker will be capable of receiving an overview of all the attendants or of some of them, just by using a device that selects the desired signals from the signals AV of the several  
20 locations and forwards them to the output audio video matrix for the following forwarding to the speaker.

Moreover, the speaker might have a graph that he is commenting to the attendants on his own screen, and these are bound to receive it full screen whilst seeing  
25 the image of the speaker himself superimposed or occupying a portion of the screen itself.

A second advantageous aspect of the present invention is that it is possible to send audio signal A coming from the speaker to an interpretation room I,  
30 wherein a simultaneous translation is carried out into



the languages required by the attendants.

The signal that is sent to each attendant therefore consists of the video signal ( $V_1, V_2, \dots, V_n$ ) *ad hoc* selected for him, to which a suitable audio  
5 signal has been associated ( $A_1, A_2, \dots, A_n$ ), therefore corresponding to the translation required by the user. It is obvious that more than one user can receive the same audio video signal AV.

Advantageously, according to the process that is  
10 herein described, it is also possible to record the audio video signal for an archive, just as it is watched by the attendants, that is with the audiovisual contributions and the television effects that have been added.

15 In so far as the apparatus apt to carry out the process so far described is concerned, within it there may substantially be envisaged a plurality of user-locations 2 (fig. 1), which are remote and/or local, and of the multimedial or interactive type, possibly  
20 equipped with a codifier/decodifier, otherwise called CODEC, with an aggregator that transforms the analog audiovideo signal AV into a digital signal, and linked to a direction room 1 that exchanges a signal AV of the analog or digital audio visual type.

25 Said signal AV contains a bunch of information relative to the conference and the speaker or the speakers that are given the right to speak from time to time, as well as other auxiliary audiovisual information.

30 Said user-locations 2 comprise audio visual

input/output means, such as for example computer or multimedial stations, tie-line linked-up locations, while the signal transmission between said locations and the direction room, and vice-versa, can take place  
5 regardless through (analog and/or ISDN) telephone lines, which can themselves be aggregate or not, satellite transmission appliances, data transmission networks (including Internet), and so on.

The signal from each remote location 2, whether it  
10 be digital or analog, is converted into an audiovideo signal, while afterwards it is sent to an audio video matrix MV1 which deals with all the signals and gives one or more output signals.

From a strictly practical point of view, direction  
15 room 1 simultaneously receives signals AV from all users 2 connected to the video conference, and it further controls the audiovideo synchronism in each single channel and, if necessary, it suitably modifies it (any possible lacks of alignment can be generated by  
20 several components: transmission, channel aggregation, reconversion).

Signals AV coming from locations 2 are each visualised by a number of monitors and they are forwarded to audiovideo matrix MV1.

25 The signals which have been selected are sent to a video mixer MIX, or computer with analogous functions, which is apt to act as an interface with a series of appliances like Personal Computers PC, Videotape recorders VD1, cameras, titlers T, audio equipment, and  
30 so on.

According to a peculiar aspect of the present invention, the employment of such video mixer MIX advantageously provides the addition to or the superimposition onto the videoconference signal, that  
5 is the signal coming from the speaker, a series of audiovisual contributions such as titles, subtitles, musical themes, soundtracks, audio and video fadings, slides and/or graphs.

Furthermore it is possible to visualise the name  
10 of the speaker that is talking in a certain definite moment, to carry out image superimpositions, to utilise and apply special effects and/or whatever other audiovisual contribution that makes the videoconference more versatile and adaptable to the needs of a specific  
15 moment.

This means that it is also possible to superimpose, back up with or create effects between the image of the speaker and films that support his talk, or graphs that he is creating himself and/or changing  
20 in that moment, and so on.

Advantageously, during a certain videoconference this makes it possible to emphasise moments of particular interest, and furthermore to underline relevant data during the talk, to highlight the aims to  
25 accomplish and/or particularly relevant news for the topic which is being dealt with.

Thus, the audio video signal which has been elaborated by the video mixer MIX or by a computer with analogous functions, is forwarded to a second  
30 audiovideo matrix MV2 and finally to a videotape

recorder VD2 which records the videoconference.

This second audiovideo matrix MV2, or visual signal sorting-out device, supplies the audio-video signals to be sent to each single user 2, whether they  
5 be remote or local.

The two input and output commutation devices of the direction room (audio video matrices MV1, MV2 or analogous devices) ensure a total compatibility between different videocommunication systems, through said  
10 plurality of CODEC or specific interfaces, so as to make it possible to carry out transmissions involving apparati with technological features that made them incompatible so far. Moreover it is possible to use just one video matrix, if this is believed necessary by  
15 the direction room, in lieu of the two abovementioned ones.

As previously properly highlighted, another peculiar feature of the present invention is given by the fact that it is possible to capture audio signal A  
20 before it reaches output audio video matrix MV2, so as to make it possible to have a simultaneous translation by one or more interpreters into the language or languages used by the attendants if these explicitly showed a need for it or if they made a clear request to  
25 the organisation.

In other words, audio signal A that is sent into interpretation room I for translation is then associated to video signal V at the output of the second audiovideo matrix MV2 in real time, in such a  
30 way that the translation or the translations are

listened to by all the attendants that requested to be supplied with such a service.

Advantageously, according to the invention, direction 1 can intervene at any moment by using  
5 audiovideo matrices, substituting audiovideo signal AV which is forwarded to one or more remote or local attendants 2 with audio video signal AVR, accomplishing an "intercom" type communication while the users who are not interested keep attending the videoconference  
10 without any disruptions or interferences.

From what explained so far follows that signal AV which is elaborated by direction room 1 must be of the analog or digital audio video type: therefore the input and output signals, i.e directed to and coming from it,  
15 which are not audio video, must be transformed before their employment and finally reftransformed at the very moment when they are to be sent to remote attendants in the analog or digital form.

These two input and output conversions at the  
20 direction room, depend on the features of the link-up with the remote users, once again categorisable as digital or analogue, which can be carried out by means that the user believes more suitable: analogue, ISDN or aggregate ISDN telephone lines, satellite transmission,  
25 computer networks (such as Internet for example), and so on.

From what described so far, it appears to be rather clear that all the attendants to the videoconference receive the audio video signal from the  
30 person that is speaking. Advantageously though, by

doubling all the incoming signals, on the speaker's screen there will be found to be shown the attendant to whom he is answering directly or with whom he intends to engage in a discussion, or in a cyclical fashion, that is all the participants to the conference (one by one or by groups, resorting to audio video multi-signal simultaneous combination devices).

To said signal which is forwarded to the speaker another signal can be added or substituted, this latter having been selected by the direction.

This is accomplished by a targeted or cyclical selection device SR, whose output signal is exclusively sent to the user that is at that moment playing the role of speaker, or otherwise to a group of users,; this is done by resorting to the second audio video matrix MV2 and whatever else is believed to be most suitable for that purpose by the direction.

It is useful to observe that a cyclical selection can take place at controllable time intervals, by dint of a timer-programmer or a computer for example.

According to another peculiar feature of the present invention, the director has the possibility of selecting the speaker who is scheduled to talk at that moment and who will be shown full screen to all the other speakers and/or attendants 2. Together with that, it is also possible to keep the audio channel of all or part of the attendants 2 active, enabling the apparatus to automatically visualise the participants that take part briefly and temporarily, in the form of windows or pointers (spots) suitably placed on the

screen.

Another extremely advantageous aspect of the present invention is the possibility of transmitting the videoconference via Internet. By suitable  
5 (aggregate or tie-line) connections between the direction room and the Internet provider, it is possible to broadcast the audio video signal AV of the videoconference, that comes from the audio-video output matrix MV2, and whatever Internet user.

10 Furthermore, by a suitable discussion group, each single user can ask questions, show examples and actively take part in the debate.

The chairperson or the person in charge of the videoconference will be capable of visualising all the  
15 communications of the final users or attendants, by a computer PCM connected to the same discussion group.

He will be capable of ascertaining whether they are worth being addressed to one of the speakers that will be then able to answer through the channels and  
20 the already described modalities of the videoconference.

If on the other hand the chairman will believe it suitable to personally let the Internet user UI contribute to the videoconference, direction room 1 is  
25 capable of carrying out an unexpected but nonetheless possible telephone link-up AV-UI, turning the Internet user UI into an actor from spectator as he was, offering him a chance to come and take part in the conference just in the same manner as that given to the  
30 other participants that are connected (provided that

said latecomer has the minimum equipment necessary for taking part in a videoconference which has the previously described modalities and features).

Advantageously, in the case of an Internet link-up, thanks to besides normal switch or ISDN telephone lines, the connection between the remote user and the provider can be carried out by dint of a mixed signal management system where the requests of the user are transmitted to the provider down the telephone lines, whereas the audio video signal of the videoconference or of the data which have been required can be received via satellite, leading to a drastic improvement of quality and increasing the speed of reception regardless of the traffic on the network and of the amount of users connected to it at that very moment.

Furthermore, using the Internet, it is possible to carry out transmission and data file exchange, regardless of the type of data therein contained, in a manner which is absolutely compatible with any type of computer or computer system.

Said remote or neighbour locations 2 may also comprise a camera and a microphone which are apt to send the audiovisual signal from a certain event like a parade or a sports match, to direction room 1 that is going to manage it in the most suitable manner.

According to the present invention, it is possible to conduct even very "intense" debates between a limited number of participants, avoiding frequent image changes; this is accomplished by subdividing the screen into adjacent windows and enabling the audio of the



entire discussion group. In this case there are found to be shown only those who are part of said restricted group of people on the screen and at the same time.

It is useful to notice that using CODECs, it is possible to control remote cameras based at locations 2. This means that the staff in the direction room is capable of showing or zooming details at their own discretion, by sending suitable directions that are bound to be executed by the camera located at the user's location.

In particular cases, it is finally possible to envisage link-ups between direction room 1 and the users exclusively via satellite.

The present invention can also be applied to other fields such as: conferences, training and refresher courses, sales, advertising, consultancy services, tourism and others.

The present invention has been described and illustrated according to one preferred embodiment, but it holds that whoever skilled in the art may well amend or change it without stepping out of the scope of the present patent.

## CLAIMS

1. Process for carrying out and managing videoconferences among remote and/or local users, characterised by the fact that it comprises the following steps:
- 5 - link-up to a direction room (1) with a plurality of both remote and neighbour locations (2), which a signal of the audio video type (AV) originates at;
  - if necessary, conversion of the audiovisual signal (AV) from each location (2), before its transfer from  
10 the place where it was generated to that where the direction room is located (1), so as to make it suitable to the type of connection and transmission which are being utilised;
  - Reconversion of the signal (AV) which has been  
15 received, if this is necessary, into the audio video format, before its arrival at the direction room (1);
  - Selection of the signal or signals to use and send away to the attendants and the speaker respectively, by an input audio video matrix (MV1);
  - 20 - Addition of the contributions and the necessary audio and/or video effects, as well as of titles, soundtracks, comments, images, graphs and so on, by a video mixer or a computer having similar functions;
  - Selection of the processed audio video signals (AV)  
25 and sending thereof to the several remote locations (2), according to the role that the users who are there located play at that moment (i.e. attendants or speakers).

2. Process according to claim 1, characterised by

the fact that while the attendants to the conference receive the audio video signal from the speaker, the speaker receives a different audio video signal which has been selected at the direction room (1).

5

3. Process according to claim 2, characterised by the fact that the speaker receives an overview of the attendants (2), or of some of them, by the employment of a targeted or cyclical selection device (SR) that  
10 selects the desired signals from the signals that arrive from the several locations, to further forward them to the output audio-video matrix (MV2), for their subsequent delivery to the speaker; said signals (AV) being capable of being simultaneously combined.

15

4. Process according to the preceding claims, characterised by the fact that the speaker is shown the graph that he is talking about to the attendants on his own screen, the attendants receiving said graph as a  
20 superimposition or within a section of the image of the speaker himself or vice-versa.

5. Process according to the preceding claims, characterised by the fact that it provides for the  
25 audio signal (A) from the speaker to be sent to the interpretation room (I) wherein a simultaneous translation into the languages required by the attendants is carried out; the signal which is sent to each attendant being therefore composed of the video  
30 signal (V1, V2, ..., Vn) selected for him, to which the

suitable audio signal (A1, A2,..., An) has been associated, i.e. the one that corresponds to the translation required by the user.

5        6. Process according to the preceding claims, characterised by the fact that more than one user can receive the same audio video signal (AV).

10       7. Process according to the preceding claims, characterised by the fact that it provides for the recording of the audio video signal for the purpose of archive or else, so well as it is actually seen by the attendants, that is enriched with the audiovisual contributions and the television effects that have been  
15 added, by a suitable videotape recorder (VD2) that receives the output signal of a video mixer (MIX) or computer with similar functions.

20       8. Apparatus for carrying out and managing videoconferences between remote and/or neighbour users, characterised by the fact that it comprises a plurality of remote and/or neighbour user-locations (2), of the interactive or multimedial type which are linked to a direction room (1) which exchanges a signal (AV) of the  
25 analog and/or digital audiovisual type with them.

9. Apparatus according to claim 8, characterised by the fact that said signal (AV) contains a series of information relative to the conference and the speaker  
30 or the speakers that are scheduled to talk, as well as

other auxiliary audiovisual information.

10. Apparatus according to claims 8 and 9, characterised by the fact that said user-locations (2) comprise audiovisual input/output means; signal transmission between said locations and the direction room, and vice-versa, taking place regardless via (aggregate or not, analog and/or ISDN) telephone lines, tie lines, satellite transmission devices, data transmission networks (including Internet), and so on.

11. Apparatus according to claims 8, 9 and 10, characterised by the fact that said remote locations (2) are equipped with analog/digital audiovisual signal conversion devices, said signal being then sent to the direction room (1) using suitable communication protocols according to the type of link which has been accomplished.

12. Apparatus according to claims 8,9,10 and 11, characterised by the fact that the direction room (1) simultaneously receives the respective signals (AV) coming from all the users (2) linked-up to the videoconference, transforms them into audiovisual signals by dint of said conversion devices and singly visualises them on a series of monitors; said signals (AV) are then channeled into an audio video matrix (MV1) that makes it possible to send just the signals coming from the speaker or speakers to the video mixer (MIX), in such a way that they are seen by all the

other attendants, with possible image fadings or other effects.

13. Apparatus according to claims 8, 9, 10, 11 and  
5 12, characterised by the fact that the signals (AV)  
selected by means of the audio video matrix (MV1) are  
forwarded to a video mixer (MIX), or a computer with  
similar functions, which is capable of interfacing with  
a number of appliances such as computers (PC), video  
10 tape recorders (VD1), cameras, titlers (T), audio  
equipment, and so on; said video mixer (MIX) making it  
possible to add to or superimpose onto the  
videoconference signal, that is to that from the  
speaker, a series of audiovisual contributions such as  
15 titles, subtitles, musical themes or soundtracks, audio  
video fadings, slides and/or graphs, visualising them  
full screen or on a portion thereof.

14. Apparatus according to claims from 8 to 13,  
20 characterised by the fact that it provides for the  
visualisation of the name of the speaker that is  
talking at a certain moment, for the carrying out of  
image superimpositions, for the use of special effects  
and/or whatever other type of audiovisual contribution  
25 that makes the conference more versatile and adaptable  
to the specific need of a certain moment; said  
apparatus further providing for the superimposition,  
the placing side by side or the creation of effects  
between the image of the speaker and of films backing  
30 up his talk, or of graphs that he himself is making or

changing at that very moment, and so on.

15. Apparatus according to claims from 8 to 14,  
characterised by the fact that the audio video signal  
5 (AV) as it is processed by the video mixer (MIX), or by  
a computer with similar functions, is sent to a second  
audio video matrix (MV2), or an analogous audiovisual  
signal sorting-out device, that provides for the signal  
to be forwarded to each single user (2), regardless of  
10 whether they be remote or local.

16. Apparatus according to claims from 8 to 15,  
characterised by the fact that the two input and output  
commutation devices of the direction room (MV1, MV2)  
15 ensure a total compatibility between different  
videocommunication systems, by said plurality of  
conversion devices, so as to provide for the  
transmission between equipments that belong to  
technological realitites that have so far been  
20 incompatible.

17. Apparatus according to claims from 8 to 16,  
characterised by the fact that the audio signal (A) is  
captured before it reaches output audiovisual matrix  
25 (MV2), so as to make it possible to carry out a  
simultaneous translation by one or more interpreters  
into the language or languages of one or more users (2)  
that may require it.

18. Apparatus according to claims from 8 to 17,  
30 characterised by the fact that the audio signal (A)

that is sent to an interpretation room (I) for the translation, is subsequently associated to the video signal (V) exiting the second audio video matrix (MV2) in real time, in such a way that the translation or the translations are respectively listened to just by all  
5 the users that make an explicit request for them.

19. Apparatus according to claims from 8 to 18, characterised by the fact that the audio video signal  
10 (AV) as elaborated by the video mixer (MIX), or by a computer with similar functions, is forwarded to a videotape recorder (VD2) that records the videoconference.

15 20. Apparatus according to claims from 8 to 19, characterised by the fact that the direction (1) can take part in whatever moment, by replacing the audio video signal (AV) which is sent to one or more attendants (2), regardless of whether they be remote or  
20 local, with an audio video signal of its own (AVR), accomplishing an "intercom" type communication while the users who are not interested keep following the videoconference without any disruption or interference.

25 21. Apparatus according to claims from 8 to 20, characterised by the fact that the signal (AV) which is elaborated by the direction room (1), is of the audio-video type: therefore the incoming signals from it that are not in the audio video format must be transformed  
30 before their utilisation and possibly retransformed



into an analog or digital form at the moment of their forwarding to remote attendants; said input and output conversions at the direction room depend on the systems used and on the analog or digital features of the link-up, with each single remote user, which be accomplished by the means that the user believes to be most suitable: analog ISDN or aggregate ISDN telephone lines, tie-lines, satellite transmissions, computer networks (e.g. Internet), and so on.

10

22. Apparatus according to claims from 8 to 21, characterised by the fact that all the attendants to the videoconference receive the audiovisual signal selected by the direction, of the person that is talking, while on the speaker's screen there is found to be visualised the attendant to whom he is answering directly, or with whom he intends to discuss, or, in a so called cyclical fashion, all the attendants to the conference (one by one or in groups); for this purpose, the doubling of all the incoming signals (AV) being provided.

15

23. Apparatus according to claim 22, characterised by the fact that said selection of the signal sent to the speaker is obtained by dint of a video matrix and a cyclical visualisation device, with the possibility of simultaneously combining more than one audiovisual sources, controlled by a timer-programmer or by a computer; the resulting signal being only sent to the speaker and/or some particular users, by the output

20

25

30

video matrix (MV2), if the direction believes it necessary.

24. Apparatus according to claims from 8 to 23,  
5 characterised by the fact that, according to schedule or else, the director can select the speaker who is scheduled to talk, who is bound to be visualised to all the other attendants to the conference and or spectators.

10

25. Apparatus according to claims from 8 to 24,  
characterised by the fact that keeping the audio channel active of all or part of the attendants to the conference (2), this makes it possible to automatically  
15 visualise the participants that take part temporarily and briefly, by the employment of windows or spots.

26. Apparatus according to claims 8 to 25,  
characterised by the fact that thanks to suitable  
20 (aggregate or tie-line) link-ups between the direction room and an Internet Provider, it is possible to transmit the audiovisual signal (AV) of the videoconference, that comes from the outgoing audio video matrix (MV2), to any Internet user.

25

27. Apparatus according to claim 26, characterised  
by the fact that by a suitable discussion group, any  
single user can ask questions, show examples and  
actively take part in the debate; a chairperson being  
30 capable of visualising on his own monitor all the

communications between the final users or spectators by a computer, and of ascertaining whether to turn them to one of the speakers that can answer using the channels and modalities of the videoconference which have  
5 already been described.

28. Apparatus according to claim 27, characterised by the fact that if the chairperson on the other hand believes it suitable to let an Internet user (UI) take  
10 part in the debate, the direction room (1) is capable of carrying out an unexpected but viable telephone link-up (AV-UI) turning the Internet user into an "actor" from being a "spectator", and offering him the possibility of getting to take part in the  
15 videoconference just in the same fashion as the other attendants who are already connected (with the proviso that the latecomer is sufficiently equipped for taking part in the videoconference with the modalities and features which were previously described).

20

29. Apparatus according to claims 8 to 28, characterised by the fact that in case of an Internet connection, besides by normal switch or ISDN telephone lines, the link-up between remote user and provider can  
25 take place thanks to a mixed signal management system where the requests made by the user are transmitted to the provider by telephone, while the audio video signal relative to the videoconference or the data which have been requested can be received via satellite,  
30 drastically augmenting the quality and the reception

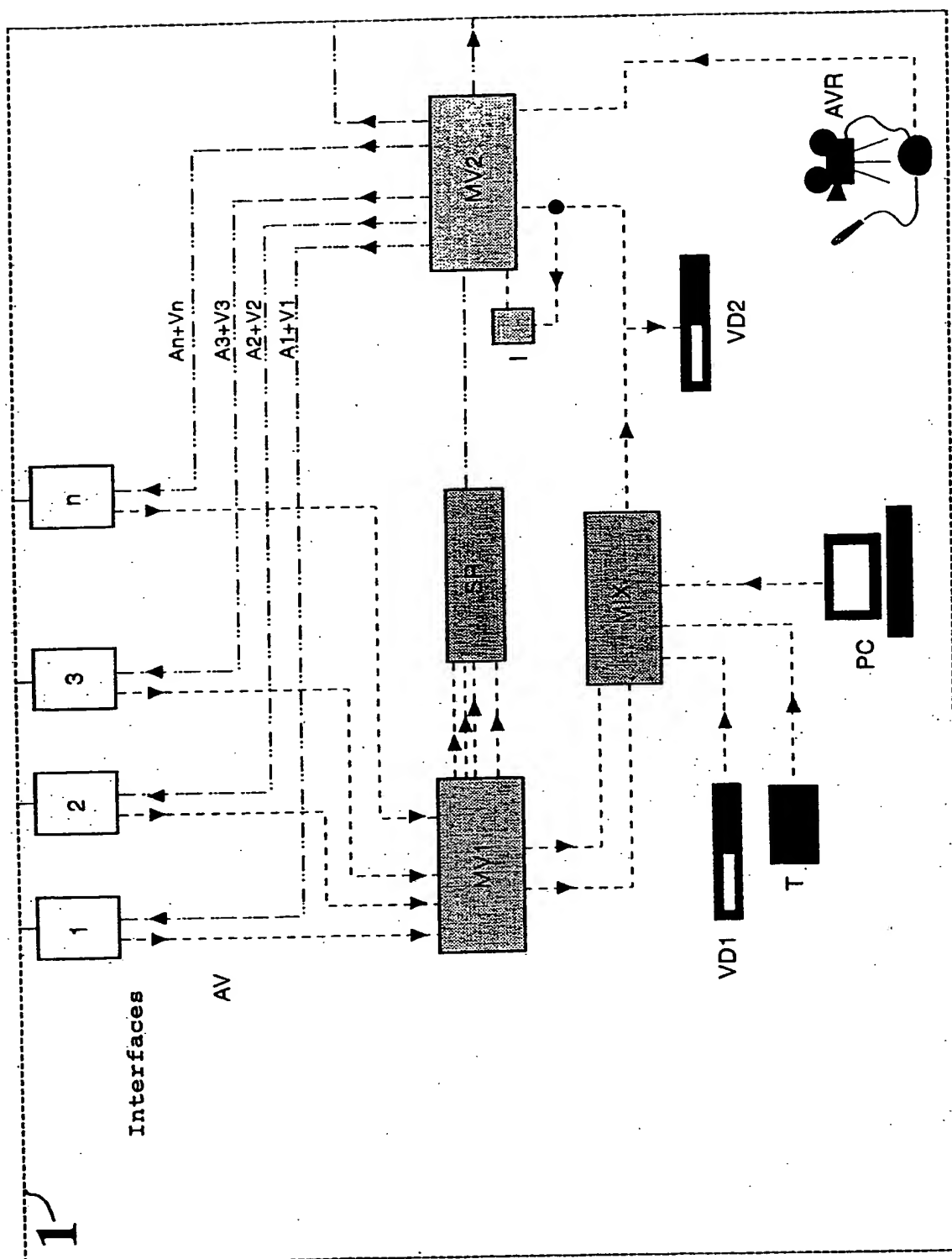
speed, regardless of the traffic on the network and of the amount of users who are connected at that moment; it being further possible to carry out the transmission and the data file exchange whatever type they are, in a manner which is absolutely compatible with whatever  
5 type of computer or computer system.

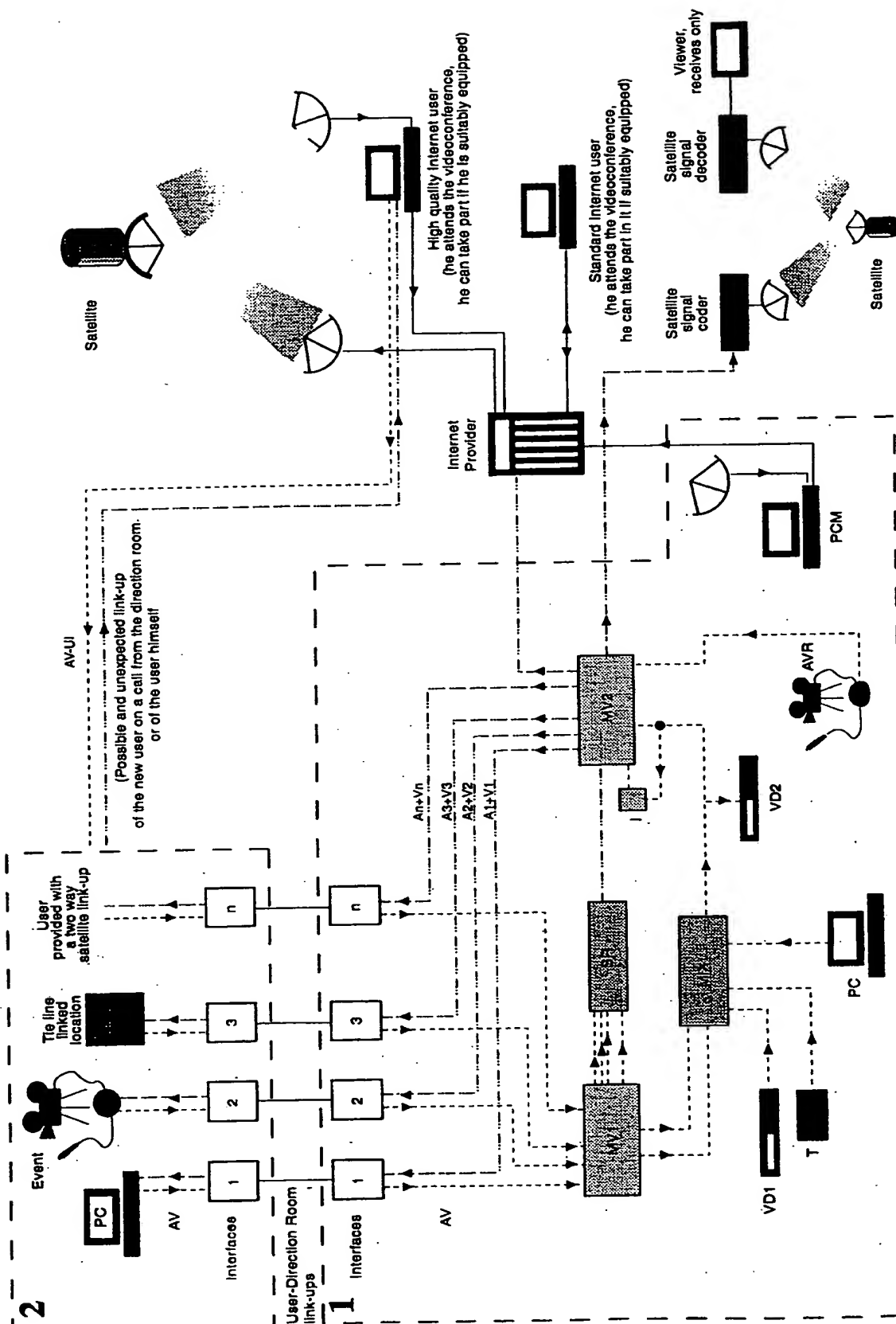
30. Apparatus according to the preceding claims, characterised by the fact that said remote or neighbour  
10 locations (2) can also comprise a camera and a microphone which are apt to forward the audiovisual signal that comes from an event, a parade, sports events or else, to the direction room (1), which is going to use it in the most suitable way.

15

31. Apparatus according to the preceding claims, characterised by the fact that the connections between the several locations, whether they be remote or local, and the direction room, are managed by dint of the  
20 normal known link-up procedures that can be by means of a telephone line carrier, by direct phone calls, by Internet network, via satellite, tie-lines, and so on.

1/2





# INTERNATIONAL SEARCH REPORT

Int'l Application No  
PCT/IT 98/00149

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 H04N7/15

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 619 679 A (FUJITSU LTD) 12 October 1994	1-3,6,8, 9,12-15, 22,23, 25,30
Y	see column 1, line 25 - column 2, line 31  see column 6, line 7 - line 12 see column 7, line 54 - column 8, line 50 see column 12, line 7 - line 44 see column 20, line 40 - column 21, line 42 see column 23, line 33 - line 35 see column 41, line 45 - column 43, line 57 see abstract; figures 1,2,11  -/-	4,5,7, 20,26-29

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

18 February 1999

Date of mailing of the international search report

05/03/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

La, V

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/IT 98/00149

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 632 659 A (AT & T CORP) 4 January 1995	4
A	see column 1, line 23 - column 2, line 15  see column 3, line 31 - line 38 see column 8, line 5 - line 42 see abstract	1,8,9, 12-14
X	WO 98 23075 A (UNISYS CORP) 28 May 1998	8-11,17, 18,21,31
Y	see page 3, line 12 - line 25	5
A	see page 6, line 11 - page 12, line 16 see abstract; figure 4	1,16
X	US 5 710 591 A (WEBER ROY P ET AL) 20 January 1998	8,19
Y	see column 3, line 19 - line 40 see column 4, line 44 - column 5, line 40 see column 6, line 32 - line 46	7
A	see abstract; figure 1	1-3,22, 23
X	KOENIG M J ET AL: "MCUs help take the worry out of being there (AT&T videoconferencing multipoint control unit)" AT&T TECHNOLOGY, WINTER 1994, USA, vol. 9, no. 4, pages 12-15, XP002093714 ISSN 0889-8979	8,16,24
A	see section "How an MCU works"	1-3,22, 23
Y	GB 2 313 251 A (MOTOROLA LTD) 19 November 1997	20
A	see page 2, line 29 - page 3, line 8  see page 4, line 33 - page 5, line 8 see page 6, line 12 - line 34 see abstract; figure 2	1,8,9, 12-14
Y	BISDIKIAN C ET AL: "MultiMedia Digital Conferencing: a web enabled multimedia teleconferencing system" IBM JOURNAL OF RESEARCH AND DEVELOPMENT, MARCH 1998, IBM, USA, vol. 42, no. 2, pages 281-298, XP000764180 ISSN 0018-8646	26-28
A	see page 282, right-hand column, line 25 - page 283, left-hand column, line 4 see page 284, right-hand column, line 32 - line 51 see page 285, left-hand column, line 33 - right-hand column, line 45 see abstract	8
	-/-	



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/IT 98/00149

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5 537 141 A (HARPER GREGORY W ET AL) 16 July 1996	29
A	see column 2, line 8 - line 13  see column 4, line 28 - line 39 see column 6, line 29 - line 36 see column 7, line 37 - line 41 see abstract	1,8,9, 12-14
A	WONG T ET AL: "Software-only video production switcher for the Internet MBone" MULTIMEDIA COMPUTING AND NETWORKING 1998, SAN JOSE, CA, USA, 26-28 JAN. 1998, vol. 3310, pages 28-39, XP002093716 ISSN 0277-786X, Proceedings of the SPIE - The International Society for Optical Engineering, 1997, SPIE-Int. Soc. Opt. Eng, USA see page 28, line 20 - line 29 see page 30, line 1 - line 6 see page 32, paragraph 2	1,8,9, 12-14
A	LUNG-SING LIANG ET AL: "DESIGN, IMPLEMENTATION AND PERFORMANCE MEASUREMENT OF MULTI-POINT VIDEO-CONFERENCE BRIDGE (MPVCB) OVER ATM SWITCH" SERVING HUMANITY THROUGH COMMUNICATIONS. SUPERCOM/ICC, NEW ORLEANS, MAY 1 - 5, 1994, vol. 3, 1 May 1994, pages 1478-1482, XP000438742 INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS see page 1478, left-hand column, line 29 - line 37 see page 1480, left-hand column, line 31 - line 39	1-3,8, 22-24

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Int'l Application No

PCT/IT 98/00149

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0619679 A	12-10-1994	JP 6296278 A US 5638114 A	21-10-1994 10-06-1997
EP 0632659 A	04-01-1995	US 5473367 A AU 682509 B AU 6594994 A CA 2122179 A,C JP 7170333 A	05-12-1995 09-10-1997 12-01-1995 31-12-1994 04-07-1995
WO 9823075 A	28-05-1998	NONE	
US 5710591 A	20-01-1998	CA 2227173 A EP 0846395 A WO 9701932 A	16-01-1997 10-06-1998 16-01-1997
GB 2313251 A	19-11-1997	NONE	
US 5537141 A	16-07-1996	AU 686795 B AU 2290795 A EP 0755608 A WO 9528804 A US 5585858 A	12-02-1998 10-11-1995 29-01-1997 26-10-1995 17-12-1996